DOCUMENT RESUME

ED 474 378 EA 032 351

| | |
|---|---|
| TITLE | Bringing Evidence-Driven Progress to Education: A Recommended Strategy for the U.S. Department of Education. Report of the Coalition for Evidence-Based Policy. |
| SPONS AGENCY | William T. Grant Foundation, New York, NY. |
| PUB DATE | 2002-11-00 |
| NOTE | 35p. |
| AVAILABLE FROM | Coalition for Evidence-Based Policy, 1301 K Street, N.W., Suite 450 West, Washington, DC 20005. Tel: 202-728-0418; Fax: 202-728-0422; e-mail: ceg@excelgov.org; Web site: http://www.excelgov.org/displayContent.asp?Keyword=prppcEvidence. For full text: http://www.excelgov.org/usermedia/images/uploads/PDFs/coalitionFinRpt.pdf. |
| PUB TYPE | Reports - Evaluative (142) |
| EDRS PRICE | EDRS Price MF01/PC02 Plus Postage. |
| DESCRIPTORS | *Educational Policy; Elementary Secondary Education; Government Publications; *Policy Formation; *Strategic Planning |
| IDENTIFIERS | *Department of Education; *Evidence Based Practice |

ABSTRACT

The recent enactment of the No Child Left Behind Act offers an opportunity to bring rapid, evidence-driven progress to U.S. elementary and secondary education. Unlike the field of human health, which has shown remarkable progress over the past 30 years, the nation's inability to raise educational achievement over the same period stands in stark contrast. To address this problem, the Coalition for Evidence-Based Policy worked collaboratively with the U.S. Department of Education to explore how the department can use its authority to advance evidence-based policy. The resulting strategy includes a department-wide effort to build the knowledge base of educational interventions proven effective through randomized controlled trials, and to provide strong incentives for the widespread use of these interventions by recipients of federal education funds. Recommendations are detailed in this report for creating an infrastructure in the department to support this strategy, building a knowledge base of effective, replicable interventions in high-priority areas, and providing strong incentives for the widespread use of proven interventions. These recommendations can be implemented within the Department's existing statutory authority, but require sustained departmental support. Three attachments provide information relating to the Coalition's work. (RT)

Reproductions supplied by EDRS are the best that can be made
from the original document.

ED 474 378

# Bringing Evidence-Driven Progress To Education:

## A Recommended Strategy for the U.S. Department of Education

**Report of the Coalition for Evidence-Based Policy**

*Advisory Board:*

Robert Boruch
Jonathan Crane
David Ellwood
Judith Gueron
Ron Haskins
Robert Hoyt
Blair Hull
David Kessler
Diane Ravitch
Laurie Robinson
Isabel Sawhill
Martin Seligman
Robert Slavin
Robert Solow
Nicholas Zill

*Executive Director:*
Jon Baron

**November 2002**

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
☑ This document has been reproduced as received from the person or organization originating it.
☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL HAS BEEN GRANTED BY

J BARON

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)



*The development of this report was funded by a grant from the William T. Grant Foundation.*

BEST COPY AVAILABLE



EA032351

## EXECUTIVE SUMMARY

The recent enactment of *No Child Left Behind*, and its central principle that federal funds should support educational activities backed by "scientifically-based research," offers an opportunity to bring rapid, evidence-driven progress – for the first time – to U.S. elementary and secondary education. Education is a field in which a vast number of interventions, such as ability grouping and grade retention, have gone in or out of fashion over time with little regard to rigorous evidence. As a result, over the past 30 years the United States has made almost no progress in raising the achievement of elementary and secondary school students, according to the National Assessment of Educational Progress, despite a 90 percent increase in real public spending per student. Our extraordinary inability to raise educational achievement stands in stark contrast to our remarkable progress in improving human health over the same time period – progress which, as discussed in this report, is largely the result of evidence-based government policies.

To address the opportunity offered by *No Child Left Behind*, the Coalition for Evidence-Based Policy undertook a collaborative initiative with the Education Department to explore how the Department can most effectively use its new authority to advance evidence-based education policy. This document – our Coalition's final report on the initiative – sets out specific recommendations for consideration by the Department leadership, as well as the broader policy community including Congress. While this report is the product of extensive discussion with Department officials and staff, its final conclusions and recommendations are those of the Coalition -- a nonprofit organization whose bipartisan Board is comprised of leading policymakers and scholars from a broad range of policy areas (attachment 2).

**Proposed Strategy: The Department should launch a major, Department-wide effort to:**

**(i) Build the knowledge base of educational interventions proven effective through randomized controlled trials – not just in small demonstration projects but in large-scale replications; and**

**(ii) Provide strong incentives for the widespread *use* of such proven interventions by recipients of federal education funds.**

The specific recommendations comprising this strategy are discussed on the next page.

**Rationale: This strategy offers a key to reversing decades of stagnation in American education, and sparking rapid, evidence-driven progress.**

Randomized controlled trials have identified a few educational interventions that are highly effective. Although rare, their very existence suggests that a concerted Department effort to build the knowledge base of these proven interventions, and spur their widespread use, could fundamentally improve the effectiveness of American education. Illustrative examples include:

- One-on-one tutoring by qualified tutors for at-risk readers in early elementary school (the 50th percentile student in the tutored group scores higher than 70-84% of the controls).
- Instruction for early readers in phonemic awareness and phonics, and guided oral reading with feedback (the 50th percentile student in each intervention scores higher than 65-70% of controls).
- High-quality preschool for low-income children (researcher-designed and implemented preschool increases percentage with high school diploma by 31% compared to controls, reduces percentage on welfare by 26% and percentage of hard-core criminals by 80%; further research is needed on how to translate these findings into broadly replicable programs).
- Life-Skills Training for junior high students (low-cost, replicable program reduces serious levels of substance abuse by over 40% by end of high school, compared to controls).

Such examples of proven effectiveness are rare – in fact, nonexistent in many areas of education – because randomized trials are uncommon in educational research and evaluation. Meanwhile, the study designs that are often used (including pre-post designs and most comparison-group designs) have been shown in careful empirical investigations to produce erroneous conclusions in many cases.

**Recommendations to create the infrastructure within the Department for this proposed strategy:**

- **The Department should identify "High Priority Areas" in which there is a critical need to (i) build the knowledge base of proven interventions or (ii) provide incentives for their widespread use.**
- **To fund the randomized trials and other recommendations below, the Department should deploy the following programs and funding sources to the maximum extent practicable:**

  (i) Office of Educational Research and Improvement funds for research and dissemination;
  (ii) "National activities" funds (for evaluation, demonstration, dissemination, and technical assistance) that the Department is authorized in law to carry out in many grant programs and areas of policy;
  (iii) A small percentage allocation from Elementary and Secondary Education Act programs, similar to that implemented by the Justice Department with informal approval of Congressional appropriators.

**Recommendations to build the knowledge base of effective, replicable interventions in High Priority Areas:**

- **The Department should focus its discretionary funds for research and evaluation on randomized trials to identify such interventions.**
- **The Department's grant programs should give applicants major incentives to use *their* discretionary funds to carry out such randomized trials.** This would enable the Department to leverage a much larger pool of resources to carry out this strategy. Such incentives would include: (i) additional funding for the applicant from the funding sources listed above; (ii) in discretionary grant programs, significant additional points in the proposal evaluation process; (iii) waiving of certain statutory and/or regulatory requirements (a strategy used effectively by the Department of Health and Human Services to get states to test welfare reform programs in randomized trials); and (iv) positive recognition and publicity.

**Recommendation to provide strong incentives for the widespread use of proven interventions:**

- **The Department's grant programs should require applicants to provide a plan for widespread implementation of research-proven interventions, with quantifiable goals.** This would apply to both formula and discretionary grant programs (discretionary programs would make this plan an important factor in the proposal evaluation process). Importantly, each applicant would be responsible for choosing which interventions, backed by randomized trials, to include in its plan.

  In High Priority Areas, the Department would require an independent evaluation, after grant award, of whether the applicant meets the goals in its plan. The Department would annually issue a high-profile report summarizing the results of these evaluations, including the progress of each state agency and other major grantee in implementing research-proven interventions.

**Conclusion**. These recommendations can all be implemented within the Department's existing statutory authority. But they will require a major, sustained commitment from the Department, involving the ongoing coordination and strategic deployment of programs and resources. Such an effort, we believe, would help end 30 years of stagnation in American elementary and secondary education, and create a new dynamic for rapid, evidence-driven progress.

## Bringing Evidence-Driven Progress To Education: A Recommended Strategy for the U.S. Department of Education

The recent enactment of *No Child Left Behind*, and its central principle that federal funds should support educational activities backed by "scientifically-based research," offers an opportunity to bring rapid, evidence-driven progress – for the first time – to U.S. elementary and secondary education. Education is a field in which a vast number of interventions, such as ability grouping and grade retention, have gone in or out of fashion over time with little regard to rigorous evidence. As a result, over the past 30 years the United States has made almost no progress in raising the achievement of elementary and secondary school students, according to the National Assessment of Educational Progress, despite a 90 percent increase in real public spending per student. Our nation's extraordinary inability to raise educational achievement stands in stark contrast to our remarkable progress in improving human health over the same time period – progress which, as discussed in this report, is largely the result of evidence-based government policies in the field of medicine.

Whether *No Child Left Behind* succeeds in bringing cumulative, evidence-driven progress to education depends critically, we believe, on how the Department of Education carries through in implementation. Thus, earlier this year the Coalition for Evidence-Based Policy and the Department of Education launched a collaborative initiative, funded by the William T. Grant Foundation, to explore how the Department can most effectively use its statutory and policy authority to advance evidence-based approaches to education policy (see attachment 1). The Coalition is a nonprofit organization, sponsored by the Council for Excellence in Government, whose bipartisan Board is comprised of leading policymakers and scholars from a broad range of policy areas (attachment 2). This document constitutes our Coalition's final report on the initiative.

In developing this report, we consulted extensively with senior Department officials and staff in the Offices of the Secretary, Deputy Secretary, Under Secretary, Educational Research and Improvement, Elementary and Secondary Education, Special Education and Rehabilitative Services, and Vocational and Adult Education. While the report reflects the valuable insights and suggestions of these individuals, its final conclusions and recommendations are those of the Coalition for Evidence-Based Policy.

### Introduction: This report sets out specific recommendations to advance evidence-based education policy, for consideration by the Education Department and broader policy community.

Our intended audience for this report is, first and foremost, the leadership of the Department of Education, and we have therefore designed all of the recommendations to be implemented by the Department within its existing statutory authority. We also offer this report and recommendations for consideration by the broader policy community, including officials and staff of the Department of Education generally, members of Congress and Congressional staff, state and local education agencies, education advocacy groups, thinktank and academic scholars, educators, and others who shape and influence the ways in which we educate our children. For major evidence-based reforms such as those outlined in this report to succeed, we believe that the larger policy community must understand why they are important, and have an opportunity to review and provide input on their design and implementation.

Our proposed strategy for the Department of Education, discussed in detail in this report, is as follows. The Department should undertake a focused and sustained effort – coordinated across Department programs and strategically deploying program, research, and evaluation funds – to:

(i) Build the knowledge base of educational interventions that have been proven effective through randomized controlled trials – not just in small demonstration projects but also when replicated on a large scale; and

(ii) Provide strong incentives for the widespread *use* of such proven, replicable interventions by recipients of federal education funds.

In this strategy, we believe, lies the key to reversing decades of stagnation in American elementary and secondary education, and bringing cumulative, evidence-driven progress – for the first time – to the U.S. educational enterprise.

This report discusses (i) why we believe that effective implementation of the "scientifically-based research" concept in *No Child Left Behind* is the key to rapid advances in the effectiveness of education policy; (ii) the challenges that the Department faces in effectively implementing this concept; and (iii) our specific recommendations to the Department for addressing those challenges and effectively advancing evidence-based education policy.

## I. Effective implementation of the "scientifically-based research" concept could fundamentally increase the effectiveness of federal education policy.

*No Child Left Behind* requires that funds in most of the major Elementary and Secondary Education Act (ESEA) programs support educational activities that are based on "scientifically-based research" (or, in some cases, "scientifically-based reading research"). ESEA programs affected by this provision include: Title I Grants to Local Educational Agencies, Reading First, Early Reading First, Even Start, Literacy Through School Libraries, Comprehensive School Reform, Improving Teacher Quality State Grants, Mathematics and Science Partnerships, English Language Acquisition State Grants, and Safe and Drug-Free Schools and Communities. "Scientifically-based research" is defined in the Act as including experimental or quasi-experimental studies, with a preference for randomized controlled trials.[1]

We believe that effective implementation of the scientifically-based research concept could have a major and enduring impact on the effectiveness of U.S. elementary and secondary education, for the following reasons.

### A. Currently, education policy is often implemented with little regard to rigorous evidence.

Federal, state, and local governments spend over $330 billion per year on public elementary and secondary education, and religious and independent schools spend an additional $30 billion. This vast enterprise encompasses thousands of educational interventions (curricula, strategies, practices, etc.), the vast majority of which have been implemented without a basis in rigorous evidence. Robert Slavin and Olatokunbo Fashola describe the problem as follows: "Change in educational practice more resembles change in fashion; hemlines go up and down according to popular tastes, not evidence. We do give lip service to research in education. Yet practices . . .

[1] *No Child Left Behind Act of 2001*, Public Law 107-110 (January 8, 2002), section 9101, paragraph 37.

go in and out of fashion, often in direct contradiction to well-established evidence, but more often in the absence of adequately rigorous and compelling research."[2]

A specific example of a widely-used practice that is contradicted by rigorous evidence is between-class ability grouping in U.S. middle and high schools, where students in a particular grade are grouped into separate classes by ability level and taught variations on the same curriculum. This type of ability grouping has become widespread despite the fact that it has little or no effect on student achievement, according to a careful review by Frederick Mosteller and others of ten randomized controlled trials of the practice. Meanwhile, other types of ability grouping that have been found effective in controlled studies – such as the Joplin Plan, which groups students across grades by ability level and uses curricula that are fitted to each group's ability – are much less common in practice.[3]

Another important example is the Drug Abuse Resistance Education (DARE) program – the nation's most widely-used school-based program to prevent drug use, operating in 75 percent of school districts nationwide, reaching an estimated 36 million children, and receiving over $200 million annually in public support. DARE's widespread adoption across the United States and in 54 other countries since its founding in 1983 occurred in the absence of any systematic evidence of its efficacy. Unfortunately, in recent years several randomized controlled trials have found that the DARE program has little or no effect on drug use among program participants.[4] Finally, after 17 years of operation, DARE announced in 2001 that it would change its curriculum in response to these studies.[5]

**B. As a result of this disconnect between practice and evidence, the U.S. has made almost no progress in raising K-12 math, reading, or science achievement over the past 30 years,** according to the National Assessment of Educational Progress, which is generally regarded as the best available measure of student achievement.[6] Remarkably, this lack of progress has occurred despite a 90 percent increase in real public spending per student in U.S. elementary and secondary schools during the same time period.[7] Our nation's lack of progress in education contrasts dramatically with our striking progress in improving health over the same time period – progress that has been driven largely by evidence-based policy in medicine, as discussed below. The contrast provides some sense of the gains that might be achieved by an evidence-based education policy.

**C. Controlled studies have identified a few highly-effective educational interventions, suggesting that if education policy were based on such findings, it could spark rapid progress.**

Randomized controlled trials have identified a few educational interventions that are highly effective. The following paragraphs provide illustrative examples. Unfortunately, the number of

---

[2] Robert E. Slavin and Olatokunbo S. Fashola, *Show Me the Evidence! Proven and Promising Programs for America's Schools* (Corwin Press, Inc., 1998), chapter 1.

[3] Frederick Mosteller, Richard J. Light, and Jason A. Sachs, "Sustained Inquiry in Education: Lessons from Skill Grouping and Class Size," *Harvard Education Review*, vol. 66, no. 4, winter 1996, pp. 797-842.

[4] Donald R. Lynam et. al., "Project DARE: No effects at 10-year follow-up," *Journal of Consulting and Clinical Psychology*, August 1999, pp. 590-593. Dennis P. Rosenbaum and Gordon S. Hanson, "Assessing the Effects of School-Based Drug Education: A Six-Year Multi-level Analysis of Project D.A.R.E.," *Journal of Research in Crime and Delinquency*, vol. 35, no. 4, November 1998, pp. 381-412.

[5] "Antidrug Program Says It Will Adopt a New Strategy," *New York Times*, February 15, 2001, p. A1.

[6] U.S. Department of Education, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), *1999 Long-Term Trend Assessment*.

[7] U.S. Department of Education, National Center for Education Statistics, *Digest of Education Statistics, 2001*.

these research-proven interventions is very small because, as discussed below, there does not now exist a systematic R&D effort, based on randomized trials, to identify and/or develop such interventions. Nevertheless, we believe these findings are very important, because by establishing that highly-effective educational interventions do exist, they suggest that a concerted Department effort to build the knowledge base of proven, replicable interventions, and spur their widespread use, could fundamentally improve the effectiveness American education.

1. **A few effective approaches to reading and math instruction have been identified in controlled studies.**

   Randomized controlled trials have consistently confirmed the effectiveness of one-on-one tutoring by highly-qualified tutors for at-risk readers in early elementary school. The effect sizes in these controlled studies generally fall between 0.5 and 1.0 standard deviations (i.e., the 50th percentile student in the tutored group scores higher than 70-84 percent of the controls), which represents a substantial effect. The effect size diminishes somewhat in the years following the tutoring intervention but remains significant.[8]

   In addition, the National Reading Panel recently undertook a comprehensive review of rigorous studies – those using randomized controls or good quasi-experimental designs – of what works in teaching children to read. The Panel's review identified several effective practices, including:

   - Teaching phonemic awareness – i.e., the ability to focus on and manipulate phonemes in spoken words (effect size averaged 0.5 standard deviations, which means that the 50th percentile student in the treatment group scored higher than about 70 percent of the controls);

   - Phonics instruction – i.e., teaching beginning readers how letters are linked to sounds and how to apply this knowledge to reading (effect size averaged 0.4 standard deviations, which means that the 50th percentile student in the treatment group scored higher than about 67 percent of the controls); and

---

[8] Batya Elbaum et. al., "How Effective Are One-to-One Tutoring Programs in Reading for Elementary Students at Risk for Reading Failure? A Meta-Analysis of the Intervention Research," *Journal of Educational Psychology*, vol. 92, no. 4, December 2000, pp. 605-619 (mean effect size for the randomized trials and matched comparison-group studies in this meta-analysis was 0.56, with larger effects for programs using higher-qualified or better-trained tutors). Barbara A. Wasik and Robert E. Slavin, "Preventing Early Reading Failure With One-To-One Tutoring: A Review of Five Programs," *Reading Research Quarterly*, vol. 28, no. 2, April/May/June 1993, pp. 178-200 (the three programs evaluated in randomized trials produced effect sizes falling mostly between 0.5 and 1.0). Barbara A. Wasik, "Volunteer Tutoring Programs in Reading: A Review," *Reading Research Quarterly*, vol. 33, no. 3, July/August/September 1998, pp. 266-292 (the two programs using well-trained volunteer tutors that were evaluated in randomized trials produced effect sizes of 0.5 to 1.0, and .50, respectively). Patricia F. Vadasy, Joseph R. Jenkins, and Kathleen Pool, "Effects of Tutoring in Phonological and Early Reading Skills on Students at Risk for Reading Disabilities, *Journal of Learning Disabilities*, vol. 33, no. 4, July/August 2000, pages 579-590 (randomized trial of a program using well-trained nonprofessional tutors showed effect sizes of 0.4 to 1.2, which were just above the median for eight other randomized trials and matched comparison-group studies of tutoring discussed in the paper ).

- Guided oral reading with feedback (effect size averaged 0.4 standard deviations, which means that the 50th percentile student in the treatment group scored higher than about 67 percent of controls).[9]

Randomized controlled trials have also identified a peer-tutoring intervention that is effective and replicable in teaching reading and mathematics to kindergarten and elementary school students. In this intervention, students are grouped in pairs, and each member of the pair alternately serves as tutor and tutee according to a well-defined protocol. At the end of each week, the class recognizes the highest-performing pairs. An important advantage of this intervention is that it is brief and inexpensive, and can be readily replicated and used in conjunction with a teacher's existing curriculum and materials. The effect size of this intervention generally falls between 0.2 and 0.6 standard deviations (which means that the 50th percentile student in the peer-tutored group scores higher than 58-73 percent of the controls).[10] This represents an educationally important effect.

2. **Controlled studies suggest that high-quality preschool can fundamentally improve life outcomes for at-risk children.**

The Perry Preschool Study, initiated in 1962, used a randomized controlled design to evaluate a program that provided high-quality preschool education to three- and four-year-old children from low socioeconomic backgrounds. The preschool curriculum emphasized active learning, in which the children engage in activities that involve making choices and problem solving, and plan their own activities, carry them out, and review them, with support and encouragement from adults.

The study followed the program participants and controls through 27 years of age, with almost no sample attrition. The study found that, by age 27, the program increased the proportion of individuals that had graduated from high school or received General Education Development certification by nearly one-third, from 54 to 71 percent, compared to the controls; quadrupled the proportion earning more than $2000 per month, from 7 to 29 percent; reduced the proportion who had ever been on welfare by a quarter, from 80 to 59 percent; and reduced the percentage of hard-core criminals (i.e., those with five or more arrests) by four-fifths, from 35 to 7 percent.[11]

A randomized controlled trial was also used to evaluate the Abecedarian Project, which provided high-quality, full-day educational child care and preschool to children from poor, uneducated families from six-weeks of age to five years, and supplementary educational support through second grade. The curriculum emphasized language and preliteracy

---

[9] "Teaching Children To Read: An Evidence-Based Assessment of the Scientific Research Literature on Reading and Its Implications for Reading Instruction," *Report of the National Reading Panel*, 2000.

[10] Patricia G. Mathes et al., "Peer-Assisted Learning Strategies for First-Grade Readers: Responding to the Needs of Diverse Learners," *Reading Research Quarterly*, vol. 33, no. 1, January/February/March 1998, pp. 62-94. Lynn S. Fuchs, Douglas Fuchs, and Kathy Karns, "Enhancing Kindergarteners' Mathematical Development: Effects of Peer-Assisted Learning Strategies," *The Elementary School Journal*, vol. 101, no. 5, May 2001, pp. 495-585. Deborah C. Simmons et. al., "Effects of Explicit Teaching on the Reading Achievement of Learning-Disabled and Low-Performing Students in Regular Classrooms," *The Elementary School Journal*, vol. 95, no. 5, May 1995, pp. 387-407.

[11] Lawrence J. Schweinhart, H.V. Barnes, and David P. Weikart, *Significant Benefits: The High/Scope Perry Preschool Study Through Age 27* (High/Scope Press, 1993). Also, an excellent summary of the main results can be found in Jonathan Crane, "Building on Success," in *Social Programs That Work*, edited by Jonathan Crane (Russell Sage Foundation, 1998), pp. 1-42.

development, particularly through adult-child transactions involving conversational language about topics of everyday interest to young children. By age 15, the program had reduced the proportion of children placed in special education by nearly half, from 48 to 25 percent, compared to controls; reduced the proportion retained in-grade from 55 to 31 percent; and improved achievement test scores by a substantial amount (0.9 standard deviations in reading and 0.6 standard deviations in math). By age 21, the program had more than doubled the proportion of individuals attending a four-year college, from 14 to 35 percent, and increased the proportion who had a good job or were in college from 40 percent to 65 percent.[12]

These study results are best characterized as early-stage research, because the projects were designed and implemented by researchers in a laboratory-like setting, and we do not yet know how to translate these findings into preschool programs that are effective and replicable on a broader scale and in a typical classroom. Nevertheless, the finding of such a powerful effect on life outcomes is very important, because it raises the real possibility that with further research we can identify broadly replicable preschool interventions that will go a long way toward addressing our nation's most pressing and intractable social problems.

3. **A Life Skills Training program for junior high school students greatly reduces the incidence of substance abuse, according to controlled studies.**

Several large-scale randomized controlled trials have demonstrated the effectiveness of Life Skills Training, a substance-abuse program for junior high school students that teaches social and self-management skills, techniques for resisting peer pressure including drug refusal skills, and consequences of drug use. The program causes very large, short-term reductions – typically in excess of 50 percent – in tobacco, alcohol, marijuana, and illicit drug use. These effects diminish somewhat over time but remain large. Long-term follow-up in one study found that, at the end of high school, the program reduced cigarette smoking by more than 20 percent; the incidence of drunkenness by 15 percent; and the proportion of students who had used tobacco, alcohol, and marijuana one or more times per month by 44 percent. The program's costs range from $5 to $10 per student per year including the cost of materials and training.[13]

Thus, randomized controlled trials have identified a few very effective educational interventions. Although these research-proven interventions are small in number, their very existence suggests that if the Department were to launch a focused and sustained R&D effort based on randomized controlled trials, it could succeed in identifying other such interventions, and over time build a sizable and highly valuable knowledge base of what works in educating American students.

**D. Pharmaceutical medicine provides an historical precedent that shows government policy can directly create a dynamic for evidence-driven advances.**

That precedent is the Food and Drug Administration's (FDA) landmark decision, in the early 1960s, to require peer-reviewed randomized controlled trials demonstrating a drug's effectiveness before allowing the drug to be marketed. The FDA decision was based on the agency's ambitious (but legally sound) interpretation of the 1962 Kefauver-Harris Amendments to the Food, Drug,

---

[12] Craig T. Ramey, Frances A. Campbell, and Clancy Blair, "Enhancing the Life Course for High-Risk Children: Results from the Abecedarian Project," in *Social Programs That Work*, op. cit., no. 11, pp. 163-183. Also, "Quality Day Care, Early, Is Tied To Achievements As An Adult," *New York Times*, October 22, 1999, p. A16.

[13] Gilbert J. Botvin with Lori Wolfgang Kantor, "Preventing Alcohol and Tobacco Use Through Life Skills Training: Theory, Methods, and Empirical Findings," *Alcohol Research and Health*, vol. 24, no. 4, 2000, pp. 250-257. Jonathan Crane, "Building on Success," in *Social Programs That Work*, op. cit., no. 11, pp. 1-42.

and Cosmetic Act, which contained a somewhat nebulous requirement for "adequate and well-controlled investigations."[14]

That FDA decision, along with parallel evidence-based policy decisions at the National Institutes of Health, have transformed life and health in the United States in ways that are often not fully appreciated. The FDA decision turned the randomized controlled trial in medicine from a rare and controversial phenomenon – which had first appeared in the medical literature only 15 years earlier – into the widely used "gold standard" for assessing the effectiveness of all new drugs and medical devices. Between 1966 and 1995, the number of clinical research articles based on randomized controlled trials surged from about 100 to 10,000 annually.[15] Since the early 1960s, FDA-required trials have demonstrated effective interventions for such serious or life-threatening conditions as hypertension, high cholesterol, hepatitis B, measles, rubella, mumps, and haemophilus influenzae type b.

The impact of these evidence-driven medical advances on the life and health in America has been truly profound. For example, measles, which during 1958-1962 afflicted an average of 503,000 people in the United States each year and caused an average of 432 U.S. deaths, and haemophilus influenzae type b, which until the 1980s was the nation's leading cause of childhood bacterial meningitis and postnatal mental retardation,[16] have both been virtually eradicated in the United States. Interventions for hypertension and high cholesterol, proven effective in FDA-required trials, helped bring about (i) a decline in the U.S. death rate from coronary heart disease, our leading case of death, by more than 50 percent since the early 1960s, and (ii) a 70 percent decline in the U.S. death rate from strokes – the nation's third leading cause of death.[17]

We believe that the Department of Education may now be in a position analogous to that of the FDA in the early 1960s. Much as the FDA's landmark decision on how to implement the Kefauver-Harris Amendments brought about rapid, evidence-driven advances in human health, so too the Department of Education's effective implementation of the "scientifically-based research" concept in *No Child Left Behind* could bring evidence-driven progress to U.S. education, and profoundly improve life outcomes for millions of Americans.

## II. Challenges facing the Department in effectively implementing the "scientifically-based research" concept.

### A. Currently, the knowledge base of educational interventions proven effective in randomized controlled trials is very small.

Perhaps the central challenge the Department faces in implementing the concept that funded activities should be based on scientifically-based research is that the knowledge base of research-proven educational interventions is very small. As discussed in the following paragraphs, this is because (i) the randomized controlled trial, which is widely regarded in other fields as the "gold

---

[14] 21 U.S.C. Section 355(d). In 1976, Congress extended this provision to cover the licensing of new medical devices (21 U.S.C. Section 360c).

[15] Mark R. Chassin, "Is Health Care Ready for Six Sigma Quality?" *The Milbank Quarterly*, vol. 76, no. 4, 1998, p. 574.

[16] "Achievements in Public Health, 1900-1999: Impact of Vaccines Universally Recommended for Children – United States, 1990-1998," *Morbidity and Mortality Weekly Report*, Centers for Disease Control and Prevention, vol. 48, no. 12, April 2, 1999, pp. 243-248.

[17] "Decline in Deaths from Heart Disease and Stroke," *Morbidity and Mortality Weekly Report*, Centers for Disease Control and Prevention, vol. 48, no. 30, August 6, 1999, p. 649-656.

standard" for assessing the effectiveness of interventions, is rarely used in educational research and evaluation; and (ii) the study designs that are commonly used – such as pre-post designs and most comparison-group designs – often produce erroneous results.

1. **Very few educational interventions currently in use have ever been tested in randomized controlled trials.**

   As Thomas Cook has noted:

   > "Very few of the major reform proposals currently on the national agenda have been subjected to experimental scrutiny. I know of no randomized evaluations of standards setting. The "effective schools" literature includes no experiments in which the supposedly effective school practices were randomly used in some schools and withheld from others. Recent studies of whole-school-reform programs and school management have included only two randomized experiments, both on James Comer's School Development Program, which means that the effects of Catholic schools, Henry Levin's Accelerated School program, or Total Quality Management have never been investigated using experimental techniques. School vouchers are a partial exception to the rule; attempts have been made to evaluate one publicly funded and three privately funded programs using randomized experiments. Charter schools, however, have yet to be subjected to this method. On smaller class size, I know of six experiments . . . . On smaller schools I know of only one randomized experiment, currently under way. In fact, most of what we know about education reforms currently depends on research methods that fall short of the technical standard used in other fields."[18]

   To this list of areas where existing practices have rarely or never been tested in randomized trials, we would add: math and science curricula, teacher and principal recruiting and training, magnet schools, character education, use of technology to improve student achievement, language instruction for limited English proficient students, vocational education, special education, dropout prevention, and after-school and summer-school programs.

2. **More generally, randomized trials are a rare phenomenon in educational research and evaluation.**

   For example, a recent review by Robert Boruch and others of the 144 contracts awarded by the Department's Planning and Evaluation Service for evaluation studies during 1995-1997 found that 51 addressed the impact of federal programs, yet only 5 of these used a randomized controlled design to measure that impact. Similarly, their review of the 84 program evaluations and other studies identified in the Department's Annual Plan for Fiscal Year 2000 found that 16 addressed the impact of federal programs, yet only one used a randomized controlled design to measure that impact.[19]

   Boruch and Christine Leow also carried out a hand search of every article in every issue of the *American Educational Research Journal* since the journal's inception in 1964, in search

---

[18] Thomas D. Cook, "Sciencephobia: Why Education Researchers Reject Randomized Experiments," *Education Next*, fall 2001, pp. 63-68.

[19] Robert Boruch, Dorothy DeMoya, and Brooke Snyder, "The Importance of Randomized Field Trials in Education and Related Areas," in *Evidence Matters: Randomized Trials in Education Research*, edited by Frederick Mosteller and Robert Boruch (Brookings Institution Press, 2002), pp. 50-79.

of articles reporting on the effectiveness of interventions in mathematics and science education. Of the 1200 articles they identified, only 35 concerned randomized controlled trials (3 percent), and there was no obvious increase in this percentage over the period 1964-1998.[20]

The rarity of the randomized controlled trial in educational research and evaluation contrasts dramatically with its pervasiveness in medical research, where as noted above approximately 10,000 articles are published annually based on randomized controlled trials.

**3. Study designs other than the randomized trial that are commonly used to evaluate educational interventions often lead to erroneous conclusions.**

Randomized controlled trials are widely considered the "gold standard" for measuring the effect of a particular intervention in medicine, psychology, welfare policy, and other areas.[21] This is because the process of randomizing subjects into either a treatment group or a control group, if properly executed (e.g., with a large enough sample size), ensures that the two groups are statistically comparable in all factors other than the intervention; therefore, the resulting difference in outcomes between the treatment and control groups can be causally attributed to the intervention.

Other study designs often lead to erroneous conclusions. For example:

**a. "Pre-post" study designs, commonly used in education research, often produce seriously biased results.**

Pre-post studies examine whether participants in an intervention improve or regress during the course of the intervention, and then attribute any such improvement or decline to the program.[22] The problem with this type of study is that, without reference to a control group, it cannot answer whether the participants' improvement or decline would have occurred anyway, even without the intervention. This often leads to erroneous conclusions about the effectiveness of the intervention.

For example, a recent randomized controlled trial of the Department of Health and Human Services' Comprehensive Child Development Program, which assigned trained case workers to connect poor families with a variety of social services through periodic home visits, found the program to be ineffective – that is, participants fared no better in all major outcomes than the members of the control group. However, both the program participants *and* the control group members experienced improvements in their lives

---

[20] Christine Leow and Robert Boruch, "Randomized Experiments on Mathematics and Science Education: Results of a Hand Search and a Machine-Based Search," report of the Campbell Collaboration and the Third International Mathematics and Science Study, University of Pennsylvania, 2000.

[21] See, for example, "The Urgent Need to Improve Health Care Quality," Consensus statement of the Institute of Medicine National Roundtable on Health Care Quality, *Journal of the American Medical Association*, vol. 280, no. 11, September 16, 1998, p. 1003. Also, Gary Burtless, "The Case for Randomized Field Trials in Economic and Policy Research," *Journal of Economic Perspectives*, vol. 9, no. 2, spring 1995, pp. 63-84.

[22] E.g., the second national evaluation of the federal Even Start program used such a study design. See *National Evaluation of the Even Start Family Literacy Program, 1994-1997 Final Report*. Washington, DC: U.S. Department of Education, Planning and Evaluation Service, 1998.

during the course of the study (e.g., a decrease in welfare dependency); thus a pre-post study design would have concluded erroneously that the program was effective.[23]

Similarly, a randomized controlled trial of the Summer Training and Education Program, a Department of Labor pilot program that provided summer remediation and work experience for disadvantaged teenagers, found that during the first summer of the program, participants lost half a grade in reading ability. However, the control group members lost a full grade in reading. Thus, a pre-post study design would have found the program's short-term impact to be negative when in fact it was positive.[24] Other examples of this phenomenon are not difficult to find.

**b. Most "comparison group" study designs, including quasi-experimental designs, also lead to erroneous conclusions in many cases.**

Comparison-group studies, which are very common in education research, compare outcomes for program participants with the outcomes of a comparison group, chosen through methods other than randomization, that did not participate in the program. For example, the comparison group may be comprised of students in neighboring classrooms or schools, or students in the same grade and socioeconomic status selected from state or national survey data. In some comparison-group studies, statistical techniques such as propensity score matching are used to create a comparison group that is statistically matched with program participants in socioeconomic and other relevant characteristics, and econometric techniques are used to control statistically for other factors that may differentially affect outcomes for program participants and comparison group members. The goal of such techniques is to simulate the results of a randomized controlled trial.

A number of careful empirical investigations have been carried out to examine whether and under what circumstances comparison-group studies can replicate the results of randomized trials. These investigations first compare participants in a particular intervention first with a control group, selected through randomization, in order to estimate the intervention's impact in a randomized study design. The investigations then compare the same intervention participants with a comparison group selected through methods other than randomization, in order to estimate the intervention's impact in a comparison-group design. The difference between the two estimates represents the bias produced by the comparison-group design.

These investigations have consistently shown that most comparison-group studies – of employment, training, and welfare-to-work programs[25] as well as K-12 education

---

[23] Robert G. St. Pierre and Jean I. Layzer, "Using Home Visits for Multiple Purposes: The Comprehensive Child Development Program," *The Future of Children*, vol. 9, no. 1, spring/summer 1999, p. 134.

[24] Jean Baldwin Grossman, "Evaluating Social Policies: Principles and U.S. Experience," *The World Bank Research Observer*, vol. 9, no. 2, July 1994, pp. 159-181.

[25] Howard S. Bloom et. al., "Can Nonexperimental Comparison Group Methods Match the Findings from a Random Assignment Evaluation of Mandatory Welfare-to-Work Programs?" MDRC Working Paper on Research Methodology, June 2002, at http://www.mdrc.org/ResearchMethodologyPprs.htm. James J. Heckman, Hidehiko Ichimura, and Petra E. Todd, "Matching As An Econometric Evaluation Estimator: Evidence from Evaluating a Job Training Programme," *Review of Economic Studies*, vol. 64, no. 4, 1997, pp. 605-654. Daniel Friedlander and Philip K. Robins, "Evaluating Program Evaluations: New Evidence on Commonly Used Nonexperimental Methods," *American Economic Review*, vol. 85, no. 4, September 1995, pp. 923-937; Thomas Fraker and Rebecca Maynard, "The Adequacy of Comparison Group Designs for Evaluations of Employment-Related Programs," *Journal of Human Resources*, vol. 22, no. 2, spring 1987, pp. 194-227; Robert J. LaLonde, "Evaluating the

programs[26] – produce very biased results, because of unobservable differences between the treatment and comparison groups.[27] In many cases, the bias is large enough that it results in erroneous overall conclusions about whether the intervention is effective, ineffective, or harmful.

These investigations are somewhat more hopeful about comparison-group designs in which the comparison group is very closely matched with the treatment group in geographic location, socioeconomic and other personal characteristics, time period in which they are studied, and methods used to collect outcome data. (Very few comparison-group studies meet this standard.) Such well-matched comparison-group designs seem to yield correct overall conclusions in most cases about whether an intervention is effective, ineffective, or harmful. However, they sometimes produce the wrong overall conclusion, and their estimates of the size of the intervention's impact usually contain a substantial amount of bias.

Comparison-group study designs have also shown important limitations in the field of medicine. A recent, well-publicized example is hormone replacement therapy for postmenopausal women. Respected, carefully-designed epidemiological studies (a type of comparison-group study) had found hormone therapy to be effective in reducing postmenopausal women's risk of heart disease and stroke. Based in part on such evidence, approximately 6 million American women currently take hormone therapy each year, at a cost of $2.75 billion annually. Over the past year, however, the results of definitive, large-scale randomized controlled trials have been released, showing that hormone therapy actually *increases* the incidence of heart attacks and stroke, as well as breast cancer. The National Institutes of Health, sponsor of one of these trials involving 16,000 women, stopped the trial in July of this year because of the findings of increased risk.[28]

Other important examples exist of medical interventions that initially appeared very effective in comparison-group (usually epidemiological) studies, but which were subsequently found in large-scale randomized trials to be ineffective or harmful. These examples include: bone-marrow transplants for women with advanced breast cancer;[29] enriched oxygen environments for premature infants;[30] beta carotene and vitamin A to

---

Econometric Evaluations of Training Programs With Experimental Data," *American Economic Review*, vol. 176, no. 4, September 1986, pp. 604-620.

[26] Elizabeth Ty Wilde and Rob Hollister, "How Close Is Close Enough? Testing Nonexperimental Estimates of Impact against Experimental Estimates of Impact with Education Test Scores as Outcomes," Institute for Research on Poverty Discussion paper, no. 1242-02, 2002, at http://www.ssc.wisc.edu/irp/. Also, Roberto Agodini and Mark Dynarski, "Are Experiments the Only Option? A Look at Dropout Prevention Programs," Mathematica Policy Research, Inc., August 2001, at http://www.mathematica-mpr.com/PDFs/redirect.asp?strSite=experonly.pdf.

[27] This literature, including the studies listed in the two preceding footnotes, is systematically reviewed in Steve Glazerman, Dan M. Levy, and David Myers, "Nonexperimental Replications of Social Experiments: A Systematic Review," Mathematica Policy Research discussion paper, no. 8813-300, September 2002.

[28] Denise Grady, "Scientists Question Hormone Therapies for Menopause Ills," *New York Times*, April 18, 2002. Gina Kolata and Melody Peterson, "Hormone Replacement Study A Shock to the Medical System," *New York Times*, July 10, 2002, p. A1. Gina Kolata, "In Public Health, Definitive Data Can Be Elusive," *New York Times*, April 23, 2002, p. F1.

[29] Gina Kolata and Kurt Eichenwald, "Health Business Thrives on Unproven Treatment, Leaving Science Behind," *New York Times*, October 2, 1999.

[30] W.A. Silverman, *Human Experimentation: A Guided Step Into the Unknown* (Oxford University Press, 1985).

prevent lung cancer;[31] dietary salt restriction to reduce hypertension;[32] and idoxuridine to treat herpes encephalitis.[33] Furthermore, systematic reviews of medical interventions that have been evaluated in both randomized trials and comparison-group studies have found that the direction of treatment effect (positive or negative) is usually, but not always, consistent across the two types of studies; however, the estimated magnitude of the effect often differs greatly, in many cases by more than two-fold.[34]

Thus, we believe there is a strong empirical basis for believing that comparison-group studies – a very common evaluation tool in education research – cannot be relied upon, by themselves, to identify effective interventions. *Well-matched* comparison-group studies do provide useful information, as discussed above; thus, we believe they can play a valuable role in a larger research agenda, serving, for example, to identify promising interventions that are then evaluated in follow-up randomized controlled trials. In medicine, for instance, carefully-designed epidemiological studies, such as the Framingham Heart Study, have been enormously valuable in identifying promising interventions to reduce heart disease and stroke – interventions which were subsequently tested and validated in randomized trials. But the evidence cautions strongly against using comparison-group studies as a final arbiter of what is effective and what is not.

**B. A second major challenge: Establishing the effectiveness of interventions not just in small demonstration projects, but also when replicated on a large scale.**

If the Department is to implement the scientifically-based research concept in way that will significantly improve educational outcomes for large numbers of American children, the interventions funded by the Department must be research-proven not only in small demonstration projects, but also when replicated on a large scale. It is possible, for instance, that an intervention may be highly effective in a small demonstration project with an unusually talented individual managing the details of implementation, but would not be effective when replicated on a statewide or national scale with other individuals managing the detailed implementation at each site. Therefore, once an intervention has been shown effective, through randomized controlled trials, in a small demonstration project, follow-up randomized trials are needed to establish the intervention's effectiveness as implemented on a larger scale, in more typical settings. Only then can we be confident that Department funds spent on that intervention, per the scientifically-based research concept, will truly improve educational outcomes.

Incidentally, this challenge of establishing the replicability of interventions is generally greater in education and other areas of social science than it is in the biological sciences. In medicine, for example, we can often be confident that a pharmaceutical drug that has been shown effective, through randomized trials, in treating a particular condition can be mass produced with the exact same chemical structure, and effectively used to treat that condition in another part of the world. Social interventions, unfortunately, cannot be replicated quite as easily, and must therefore be rigorously evaluated for effectiveness when implemented on a large scale.

---

[31] Charles Marwick, "Trials Reveal No Benefit, Possible Harm Of Beta Carotene and Vitamin A for Lung Cancer Prevention," *Journal of the American Medical Association*, vol. 275, no. 6, February 14, 1996, p. 422.

[32] J. Swales, "Population Advice on Salt Restriction: The Social Issues," *American Journal of Hypertension*, vol. 1, part 1, January 13, 2000, pp. 2-7.

[33] Charles A. Alford, Jr. and Richard J. Whitley, "Treatment of Infections due to Herpesvirus in Humans: A Critical Review of the State of the Art," *Journal of Infectious Diseases*, vol. 133, supplement, June 1976, pp. A101-A108.

[34] John P.A. Ioannidis et. al, "Comparison of Evidence of Treatment Effects in Randomized and Nonrandomized Studies," *Journal of the American Medical Association*, vol. 286, no. 7, August 15, 2001, pp. 821-830. This paper also summarizes the findings of earlier reviews that have been carried out on this topic.

**C. A third major challenge: How to induce grant recipients to implement research-proven interventions.**

Even if a large knowledge base of proven, replicable interventions existed, the Department would still face the challenge of how to get grant recipients to implement them effectively, per the statutory requirement that funded activities shall be "based on scientifically-based research."

The challenge is best illustrated by the experience of a precursor to the "scientifically-based research" concept, contained in the Comprehensive School Reform Demonstration (CSRD) program. The CSRD program, enacted in 1997 as part of Public Law 105-78, contained a legislative requirement that grant recipients adopt comprehensive school reforms that are based on "rigorous research and effective practices." The legislation identified 17 examples of such school reforms for grant recipients to consider, but also gave recipients the option to adopt another reform approach "based on rigorous research."[35] The Department's guidance for program participants incorporated this requirement and also provided explanatory material about the components of reliable research and effective practices.[36] In addition, in 1999 the American Institutes for Research (AIR) – an independent, internationally-recognized research organization – published a user-friendly guide for educators, modeled on *Consumer Reports*, that rated various school reform models on the strength of evidence supporting their effectiveness.[37]

Effective implementation of CSRD has been hampered, first, by the inadequate knowledge base issue cited above – specifically, the dearth of school reform models proven effective in randomized trials. Of the many comprehensive school reform models in existence, including the 17 cited in the legislation, only James Comer's School Development Program and Robert Slavin's Success for All have been subjected to a randomized controlled trial. The School Development Program was found to be modestly effective in one site and ineffective in another[38]; the Success for All trial is ongoing and results are not yet available.

But even if we sidestep that issue and assume a lesser standard of evidence than the randomized trial – which is the approach AIR took in rating the various school reform models – it's not clear that CSRD has succeeded even in getting grant recipients to adopt school reforms meeting this lower standard.

For example, in CSRD's first year of operation (1998), 29 percent of schools receiving CSRD funds adopted school reform models that AIR rated as having "strong" or "promising" evidence of effectiveness (based on comparison-group and other nonrandomized studies). The hope of many CSRD supporters was that this percentage would increase as the program matured, consistent with CSRD's goal of supporting proven reform models. In fact, just the opposite seems to have occurred. Preliminary data from a survey of 19 states, shown in attachment 3, indicates that in 2001-2002 just 18 percent of schools receiving CSRD funds adopted school reform models rated by AIR as having a "strong" or "promising" research base. As shown in the attachment, CSRD funds are increasingly going toward school reform models whose research base is rated by AIR as "marginal" and to models not reviewed by AIR (e.g., because they are not comprehensive programs or have no research backing).

---

[35] House Report 105-390, November 7, 1997, p. 97.

[36] *Guidance on the Comprehensive School Reform Demonstration Program*, U.S. Department of Education, August 11, 2000.

[37] Rebecca Herman et. al., *An Educators' Guide to Schoolwide Reform*, American Institutes for Research, 1999.

[38] Thomas D. Cook, H. David Hunt, and Robert F. Murphy, "Comer's School Development Program in Chicago: A Theory-Based Evaluation," *American Educational Journal*, vol. 36, no. 3, fall 1999, pp. 543-597.

Thus, a central challenge the Department faces in implementing the "scientifically-based research" requirement is that the mere inclusion of this requirement in statute and program guidance does not seem to guarantee that grant recipients will actually implement research-based interventions. Nor does the existence of *Consumer Reports*-type ratings of interventions based on evidence of effectiveness guarantee such implementation. Further incentives or inducements, we believe, are needed, and therefore comprise a central element of our recommendations below.

## III. Recommendations: The Department should launch a major, concerted effort to (i) build the knowledge base of proven interventions and (ii) spur their widespread use.

Our proposed Department strategy is as follows. The Department should launch a focused and sustained effort – coordinated across Department programs and strategically deploying program, research, and evaluation funds – to:

(i) Build the knowledge base of educational interventions that have been proven effective through randomized controlled trials – not just in small demonstration projects but also when replicated on a large scale; and

(ii) Provide strong incentives for the widespread *use* of such proven, replicable interventions by recipients of federal education funds.

This proposed strategy is comprised of 11 specific recommendations, discussed in this section. They include: (i) our recommendations for creating the infrastructure within the Department to carry out this strategy; (ii) our recommendations for building the knowledge base of effective, replicable interventions; and (iii) our recommendations for spurring widespread use of these proven interventions. The final recommendation outlines a process whereby evidence-based reforms in the Department's grant-making procedures, including those proposed in this report, would themselves be rigorously evaluated in randomized controlled trials.

The primary audience for these recommendations, as noted earlier, is the leadership of the Department of Education, and accordingly all of the recommendations are designed to be implemented by the Department in the near term within the Department's existing statutory authority. We believe that these recommendations could readily be adapted and applied to other major components of U.S. pre-K-12 education policy, such as the National Science Foundation's funding of research on science and mathematics instruction in elementary and secondary schools, the Department of Health and Human Services' Head Start program, and major state-funded pre-K-12 education programs. Such an undertaking is beyond the scope of this project, but we believe may be worth pursuing in future initiatives.

### Recommendations to create the infrastructure within the Department to carry out this effort.

**1. The Department should establish a Committee, chaired by the Under Secretary or above, to coordinate the implementation of this effort across the Department.**

We believe such a Coordinating Committee is needed because this effort, as discussed below, will require coordinated decision-making and deployment of resources across the Department's program, research, and evaluation activities. Without specifying the structure and composition of the Coordinating Committee, we believe that:

- It should include representatives from the Department's major program, research, and evaluation activities (e.g., the Assistant Secretaries);

- It should be charged with coordinating the Department-wide implementation of this effort, including the recommendations in this report; and

- To ensure that the Committee has the authority to effectively advance a Department-wide effort of this magnitude: (i) it should be chaired by the Under Secretary (or above); and (ii) when the Department's non-research grant programs seek Department approval for budgets, contracts, and *Federal Register* notices for research and evaluation, they should be required, as part of the approval process, to obtain the concurrence of the Committee.

2. **The Department should identify "High Priority Areas" in which there is a critical need to (i) build the knowledge base of effective interventions or (ii) spur widespread use of such interventions.**

   a. **First, we believe the Department should identify, and publish a list of, High Priority Areas for building the knowledge base of interventions proven effective** and replicable in randomized controlled trials. The Department would develop this list of High Priority Areas based on such factors as: (i) the likelihood of identifying effective interventions in a given area that would substantially improve educational and life outcomes of participants; (ii) a large federal investment of funds in the area, the impact of which could be greatly increased if effective interventions are identified; and (iii) the degree to which the Department has discretion over how program funds are spent in the area, and thus the ability to focus those funds on randomized trials.

      As illustrative examples:

      - The Department might choose to designate Comprehensive School Reform as a High Priority Area if it believes, based on existing comparison-group studies, that (i) randomized trials are likely to identify effective, replicable school reform models, and (ii) such models could greatly improve educational outcomes (reading and math achievement, special education referrals, high school graduation rates).

      - The Department might choose family literacy as a High Priority Area if it believes that (i) randomized trials are likely to identify effective interventions in this area; and (ii) such interventions could increase the impact of the large, ongoing federal investment in the Even Start Family Literacy Program – a program which currently does not appear to be effective.[39]

      - The Department might choose character education as a High Priority Area, because the Department directly awards grants in this area through a discretionary (i.e., competitive) granting process. Thus, the Department can use the proposal evaluation process to provide strong incentives for applicants to structure their activities as randomized trials (as discussed under recommendation 8c, below).

   b. **Second, we believe the Department should identify, and publish a list of, High Priority Areas for incentivizing the widespread *use* of interventions proven effective** and replicable

---

[39] Per the results of a randomized controlled evaluation of Even Start, reported in Robert G. St. Pierre et. al., "Improving Family Literacy: Findings From the National Even Start Evaluation," Abt Associates, September 1996. A second randomized controlled trial of the Even Start program is underway, with a scheduled completion date in the summer of 2003.

in randomized trials. The Department would develop this list of High Priority Areas based on such factors as (i) the existence of interventions in a given area that have been proven effective and replicable in randomized trials, and (ii) the potential impact of widespread replication of such interventions on educational and life outcomes. For example, the Department might choose to identify early reading as a High Priority Area, based on (i) the existence of early reading interventions, discussed earlier, that have been shown effective through randomized trials in raising reading achievement; and (ii) a belief that early reading success may be closely tied to later school achievement and, ultimately, successful life outcomes.

These lists would not represent the Department's priority areas in the large sense – e.g., for the purposes of the Department's budget request to Congress. Rather, they would serve the more limited function of identifying areas in which the Department believes that additional investment in building the knowledge base of, or disseminating, proven interventions will be most productive.

3. **Having identified High Priority Areas, the Department should estimate the funding needed to build the knowledge base of proven interventions and spur their widespread use in these Areas**, per the recommendations outlined in this report.

The Department-wide effort outlined in this report – particularly the randomized controlled trials to identify effective interventions in High Priority Areas – will require a significant commitment of resources (which we believe is minor, however, in relation to the $300+ billion in annual education spending whose effectiveness it seeks to improve). To give some idea of the cost of such trials:

- Two of the infrequent cases where the Department has funded randomized controlled trials in recent years include the School Dropout Demonstration Assistance program and the Upward Bound program. The Dropout Demonstration study involved randomized trials in each of 16 sites, and cost $7.3 million over 1991-1995. The Upward Bound study involved randomized trials in each of 67 sites, and cost over $5.4 million during 1992-1996.[40]

- Tennessee's Student-Teacher Achievement Ratio (STAR) Project was a large-scale randomized controlled trial, funded by the state of Tennessee, that examined the effect of reducing class size in early elementary school on student achievement. The trial, involving 79 schools and over 11,000 students, cost $12 million over 1985-1989.

- Success For All, a comprehensive school reform program, is currently being evaluated in a randomized controlled trial, funded by the Department of Education. The trial, involving 60 schools over a five-year period, is expected to cost over $6 million. The Department estimates that a larger initiative to evaluate 5 to 10 school reform models in similar trials would cost $42 million over a six-year period.

The Department should use data such as this to estimate the funding required to carry out the recommendations below in the High Priority Areas that the Department has identified.

4. **To obtain the required funding, the Department should strategically deploy existing programs and funding sources in support of this effort.**

What follows is a list of possible programs and sources of funding that could be deployed for this effort. Deployment of these resources would presumably involve a collaboration between the Coordinating Committee (recommendation 1) and the particular program office(s) involved. We also

[40] Robert Boruch, Dorothy De Moya, and Brooke Synder, op. cit., no. 19, p. 57.

show the funding levels associated with the listed programs and funding sources, with the important caveat that in most cases only a portion of these funds could be deployed for this effort since the funds support other important activities as well.

a. **Office of Educational Research and Improvement (OERI) funds for research and dissemination** ($122 million in FY 2002).

b. **"National activities" funds within various Department grant programs.** Specifically, the authorizing language for many Department grant programs allows the Department to carry out "national activities" with a small amount of program funds – activities that typically may include: the evaluation of program as a whole as well as of specific interventions that it funds; dissemination of best practices among grant recipients; funding of demonstration projects; and technical assistance for grant recipients.

   There are often many demands on these national activities funds, because they are one of the few sources of funds for program offices and the Office of the Secretary to engage in discretionary activities. Program offices typically prefer to spend these funds primarily on technical assistance for grant recipients rather than on evaluation activities such as those we outline in the recommendations below. We estimate that in FY 2002, for example, program offices will spend only about one-third of their national activities funds on evaluation, and two-thirds on other activities that primarily include technical assistance.[41]

   However, Administrations can, and often do, deploy the national activities funds so as to advance policy priorities. Consistent with our view that the Department, and the Administration, should embrace the strategy outlined in this report as a central priority, we recommend that the Department focus these funds, to the maximum extent practicable, on the randomized trials and other activities in the recommendations that follow. The Coordinating Committee (recommendation 1) could be used to focus the national activities funds in this way, through its role in reviewing the grant programs' plans for spending these funds on research and evaluation.

   The following is a description of the national activities authorized for the larger Department grant programs:

   - Reading First. The Department may use up to 2.5 percent of program funds to carry out national evaluation and technical assistance activities. 2.5 percent amounts to $22.5 million in FY 2002.

   - Even Start. The Department may use up to 3 percent of program funds to carry out national evaluation, technical assistance, and related activities. 3 percent amounts to $7.5 million in FY 2002.

   - Safe and Drug-Free Schools. Congress appropriated $135 million in FY 2002 for national demonstration, evaluation, dissemination, and technical assistance activities, as well as an independent impact evaluation of the program.

[41] This estimate is based on data for the following programs, each of which has a small set-aside for national activities: Reading First, Even Start, Comprehensive School Reform, Teacher Quality State Grants, Educational Technology State Grants, 21st Century Community Learning Centers, Public Charter Schools, and Magnet Schools Assistance.

- Indian Education. Congress appropriated $3.2 million in FY 2002 for national research, evaluation, and other activities in this area.

- Early Reading First. The Department may use a total of $3 million in program funds over FY 2002-2005 for national evaluation activities.

- Comprehensive School Reform. The Department may use up to 1 percent of program funds for national evaluation activities. 1 percent amounts to $2.4 million in FY 2002.

- Educational Technology Grants. The Department may use up to 2 percent of program funds for national evaluation, dissemination, and technical assistance activities. 2 percent amounts to $14 million in FY 2002.

- 21st Century Community Learning Centers. The Department may use up to 1 percent of program funds for national evaluation and technical assistance activities. 1 percent amounts to $10 million in FY 2002.

- Public Charter School Grants. The Department is required to use $8 million in program funds in FY 2002 for national evaluation, dissemination, and technical assistance activities.

- Magnet Schools Assistance. The Department may use up to 2 percent of program funds for national evaluation, dissemination, and technical assistance activities. 2 percent amounts to $2.2 million in FY 2002.

c. **National activities authorized elsewhere in law**, as follows:

- National Assessment of Title I. Congress appropriated $9 million in FY 2002 for (i) a national evaluation of programs and interventions funded under Title I of ESEA; (ii) a longitudinal study of schools receiving assistance through Title I Grants to Local Educational Agencies; and (iii) national demonstrations of innovative practices.

- Individuals with Disabilities Education Act. Congress appropriated $78 million in FY 2002 for national research and demonstration activities to produce "new knowledge" about how to improve educational outcomes for children with disabilities, and to translate that knowledge into practice. Congress earmarked approximately $8 million of these funds for specific projects, leaving the Department with discretionary use over the remaining $70 million.

- Carl D. Perkins Vocational and Technical Education Act. Congress appropriated $12 million in FY 2002 for national research, evaluation, demonstration, and dissemination activities in this area.

- Fund for the Improvement of Education. Congress established this Fund for the Department to carry out national research, evaluation, and demonstration activities to improve the quality of elementary and secondary education. Most of the money appropriated for the Fund in FY 2002 has been earmarked by Congress for specific projects; however, the Department retains discretionary use over $12 - $15 million.

- Title IX Evaluations. Some ESEA programs do not have specific provisions in their authorizing language allowing them to carry out national evaluation activities. Title IX of ESEA allows certain of these programs to use up to 0.5 percent of their funds for such activities. The largest ESEA program to which this provision applies is Improving Teacher

Quality State Grants, where it yields approximately $14 million in funding for national evaluation activities.

d. **A small percentage allocation from ESEA programs.** If the Departmental funds available from the above sources are not sufficient to carry out the strategy outlined in this report, we recommend that the Department seek a small percentage allocation from ESEA program funds to generate the needed funding. An important advantage to the Department of carrying out such a percentage allocation is that it could generate a significant pool of funds that would not be encumbered by requirements that the funds be used for particular programs or activities (unlike the national activities funds, discussed above). In FY 2002, a 2 percent allocation from ESEA programs would yield approximately $400 million.

Although the Department has the statutory authority to transfer limited amounts of funds in this way without obtaining Congressional approval,[42] we recommend that the Department not use this authority, since doing so may produce an unfavorable Congressional reaction. Rather, if the Department decides to pursue this approach, we recommend that it:

- Seek the informal approval of the Congressional Appropriations Committees to implement a small percentage allocation of this type. Such an approach has been successfully carried out by the Department of Justice. Specifically, since FY 1994 the Justice Department has transferred 1 to 3 percent of funds from certain crime prevention programs to the National Institute of Justice, to carry out research and evaluation. This percentage allocation is carried out each year with the informal approval of the Congressional Appropriations Committees, and has endured through changes in Administrations and in the parties controlling Congress. It has, in some years, yielded more than $10 million.

- Over the longer-term, seek a statutory change authorizing such a percentage allocation. A successful precedent for this approach is found in Section 241 of the Public Heath Service Act, which authorizes the Department of Health and Human Services to allocate 1.25 percent of funding from certain public health service programs, to fund program evaluation activities. This authority, first established in 1970, will yield nearly $400 million in FY 2003.

5. **The Department should establish a process to qualify organizations with the capability to carry out high-quality randomized controlled trials and other evaluations.**

Specifically, we recommend that the Department establish:

a. **A process to qualify "Impact Evaluators"** – i.e., organizations that have the capability to work with state/local agencies and educators to carry out high-quality randomized controlled trials; and

b. **A process to qualify "Process Evaluators"** – i.e., organizations that have the capability to work with grant recipients to carry out high-quality evaluations of (i) the extent to which a grant recipient is implementing proven interventions in its educational activities, and (ii) the fidelity of such implementation.

We recommend that the Department structure these as streamlined, open processes designed to qualify as many organizations as possible with the appropriate expertise. The Department may also wish to engage an independent panel of experts to review the applications and make the qualification

[42] Section 304 of the FY 2002 Appropriations Act for the Departments of Labor, Health and Human Services, Education, and Related Agencies, Public Law 107-116 (January 10, 2002).

decisions. The Department would publish and regularly update a list of the qualified Impact Evaluators and Process Evaluators.

**Recommendations to build the knowledge base of interventions proven effective and replicable in randomized trials.**

6. **OERI should sponsor "design competitions" to fund the design of new educational interventions in High Priority Areas, as well as their evaluation in randomized trials.** Recently, OERI has taken important steps in this direction. For example, in December OERI initiated such a design competition to identify effective, replicable preschool curricula. The competition will fund 10 randomized controlled trials of preschool curricula over a four-year period, each involving at least 10 classrooms with a total of 150 children. OERI is also funding an independent evaluation of trial outcomes.

   Such OERI-sponsored design competitions, evaluated through randomized controlled trials, could be used to help build the knowledge base of effective interventions in a number of areas, such as: elementary school math curricula, character education, and reading and math tutoring programs using volunteer tutors. However, for an intervention to be demonstrated effective *and* broadly replicable, the Department will need to sponsor follow-up randomized trials for those interventions initially found effective in the design competitions, to establish their replicability in various settings and populations.

   The cost of such design competitions and follow-up trials is large in relation to OERI's overall research budget ($122 million in FY 2002), from which OERI must also fund other important research activities (e.g., basic research on the cognitive processes by which children learn). For example, the projected four-year cost of OERI's preschool curriculum competition is $16 million, not including the independent evaluation. The Department may therefore wish to provide supplementary funds for these design competitions from funding sources listed under recommendation 4, and/or use other mechanisms discussed below to carry out the follow-up randomized trials.

7. **Program evaluations carried out by the Department with its discretionary funds should focus, to the maximum extent practicable, on randomized trials in the High Priority Areas.**

   The Department will spend approximately $64 million in FY 2002 on the evaluation of Department programs, primarily by drawing on the national activities funds. In previous years, randomized controlled trials have comprised a very small fraction of these evaluations, as discussed earlier. Recently, however, the Department's current leadership has begun to shift these evaluations toward randomized controlled designs. In FY 2002, for example, the Department will spend approximately 39 percent of these evaluation funds on randomized controlled trials; in FY 2003, the Department proposes to spend approximately 61 percent of these funds on randomized trials.

   We strongly endorse this trend. If the Department can couple this shift toward randomized designs within its evaluation activities with a major, Department-wide redeployment of resources toward such activities (per recommendation 4 above), we believe this will represent a major step forward in building the knowledge base of proven, replicable interventions. We believe these randomized controlled evaluations should focus not just on assessing the overall effectiveness of Department programs, but also, as discussed earlier, on evaluating specific interventions within each program with an eye toward identifying effective, replicable interventions in the High Priority Areas.

8. **The Department's formula and discretionary grant programs should give applicants major incentives to structure some of their activities as randomized trials in the High Priority Areas.** We believe that, by providing such incentives, the Department can effectively leverage a much larger pool of resources than its own discretionary funds, and thereby greatly increase the size and scope of the effort to build the knowledge base of proven, replicable interventions.

   Specifically, we recommend that in the Department's discretionary and formula grant programs:

   a. **Grant applicants should be given significant incentives (detailed under b, below), to structure some of their proposed activities as randomized trials that –**

      i. **Test interventions in the Department's High Priority Areas**, not just in demonstration projects but also when replicated on a large scale;

      ii. **Use sound study designs** that minimize sample attrition and include outcome measures that are reliable and valid; and

      iii. **In appropriate cases, are independently evaluated.** In such cases, the Department would ask applicants to team with an Impact Evaluator in the design of the trial and measurement of the trial's outcomes. Independent evaluation would be appropriate in cases where the trial has particularly important, near-term policy implications, necessitating an especially high level of evaluation rigor and integrity (e.g., when the applicant is testing a well-developed and replicated intervention that, if proven effective, would be ready for large-scale implementation).

         Even when a randomized trial will not be independently evaluated, the applicant may decide to team with an Impact Evaluator to obtain the expertise needed to design and implement it.

      Because the randomized design must be built into the project from its inception, entailing up-front costs, applicants will need to address in their proposals how they plan to make funds available for the trial in advance of the school year when it will be implemented (e.g., by saving a portion of current-year funds to carry out the randomized trial in the next school year).

   b. **The incentives that these grant programs would give applicants to carry out such randomized trials could include one or more of the following:**

      i. **Additional funding for the applicant's activities**, from funding sources identified per recommendation 4, above.

      ii. **Funding to design the trial and carry out the evaluation of its outcomes**, from funding sources identified per recommendation 4. (In projects involving an independent evaluation, the Department would provide such funding directly to the Impact Evaluator to ensure the evaluation's independence.)

      iii. **In discretionary grant programs, significant additional points in the proposal evaluation process** for applicants that structure some of their proposed activities as randomized trials in High Priority Areas.

      iv. **Waiving statutory and/or regulatory requirements for applicants that propose such randomized trials, and providing them with other flexibility** in their use of grant funds.

Federal welfare policy provides an important and successful precedent for this approach. Specifically, in the mid-1980s, the Department of Health and Human Services (HHS) implemented a demonstration waiver policy, under which HHS waived certain provisions of federal welfare law to allow state-level grantees to test new welfare reform approaches, but only if the grantees agreed to evaluate their reforms in randomized controlled trials. This policy directly resulted in more than 20 large-scale randomized controlled trials of welfare reform programs from the mid-80s through the mid-90s – studies which taught us much of what we know today about what works in moving people from welfare to work.

Importantly, HHS's policy of linking waivers to a requirement for randomized trials was not specifically authorized or directed in statute, but rather was developed and implemented administratively, through a creative application of HHS's statutory authority. The Department of Education has similar waiver and flexibility authority, which it could also tie to a requirement for randomized trials. Specifically:

- **Under the Department's "waiver" authority**, the Secretary may waive certain statutory or regulatory requirements of ESEA for state/local education agencies that submit a waiver request to the Department. We recommend that the Department (i) officially encourage state/local agencies to request waivers from the Department to carry out randomized trials in the High Priority Areas; and (ii) implement an expedited process for approving such waiver requests.

- **Under the Department's State-Flex and Local-Flex programs**, the Department currently allows state/local educational agencies that have been selected through a competitive process to consolidate certain federal education funds and use those funds for any educational purpose authorized under ESEA. We recommend that, in these two programs, the Department provide significant additional points in the proposal evaluation process for applications that include randomized trials in High Priority Areas.

- **Under the Department's "Consolidated State Applications" process**, a state may submit a single, consolidated application to the Department to obtain funds under many ESEA formula grant programs, rather than submit separate applications for each program. We recommend that the Department require states, in their consolidated applications, to include randomized trials in High Priority Areas.

v. **Positive recognition and publicity**. For example, large urban school districts that are under significant pressure to improve their performance may be favorably inclined to propose and participate in randomized trials if doing so will bring positive recognition from the Department.

c. **Specific examples of how the Department can use these incentives to get applicants to propose randomized trials in High Priority Areas:**

- **In the Department's programs that include a significant discretionary component (i.e., where the Department makes some or all grants through a competitive process)**, the Department could give applicants one or more of the following incentives to propose such randomized trials : (i) additional points in the evaluation process; (ii) additional funding for the applicant's activities and/or for the design and evaluation of the trial; (iii) waivers of statutory or regulatory requirements that might otherwise hinder or slow the implementation of the trial; (iv) positive recognition and publicity.

Examples of Department programs that include a significant discretionary component and could deploy the incentives in this way include: Early Reading First, Safe and Drug-Free Schools, Special Education, and Character Education.

The Department has already taken an important first step toward implementing this approach in the Character Education program. Specifically, the program's process for evaluating grant applications, as implemented this spring, provides 20 additional points for projects that employ an experimental design with random assignment, or if not feasible, a quasi-experimental design with carefully matched comparison conditions (100 points are otherwise available).

- **In Department programs that make formula grants to the states, which then make competitive grants to local entities** (e.g. Reading First, Comprehensive School Reform, Safe and Drug-Free Schools), the Department could give incentives to the states to provide additional points in the *state-level* proposal evaluation, and/or additional funding, to local applicants that propose randomized trials. Such incentives for the states could include additional funding, waivers of statutory/regulatory requirements, and/or positive recognition and publicity.

- **In the Department's formula grant programs that allow states to reserve a portion of their grant funds for state-level discretionary activities** (e.g., evaluation, demonstration, technical assistance, and other activities), the Department can give incentives to the states to spend these discretionary funds on randomized trials in High Priority Areas. Such incentives for the states could include additional funding; waivers of statutory/regulatory requirements, and/or positive recognition and publicity.

The Department could apply this approach in most of its large formula grant programs, which usually allow states to reserve a significant portion of their grant funds for discretionary activities. For example:

Under the Reading First, each state may reserve up to 20 percent of its grant funds for discretionary activities such as providing professional development to K-3 teachers, strengthening higher education programs that train K-3 teachers, and assessment and evaluation of local agencies receiving subgrants under the program. 20 percent amounts to approximately $170 million total (i.e., for all 50 states) in FY 2002.

Under Improving Teacher Quality State Grants, each state is required to spend 2.5 percent of its grant funds for discretionary activities such as providing teacher mentoring, providing professional development to teachers, and developing merit-based performance systems for teachers. 2.5 percent amounts to approximately $70 million total (i.e., for all 50 states) in FY 2002.

Under Educational Technology Grants, each state may use up to 5 percent of its grant funds for discretionary activities such as developing innovative strategies for using technology to deliver courses and curricula, and measuring the effectiveness of local activities funded by the program. 5 percent amounts to approximately $30 million total (i.e., for all 50 states) in FY 2002.

Under State Grants for Innovative Programs, each state may use up to 15 percent of its grant funds for discretionary activities such as monitoring and evaluation of local activities funded

by the program, school improvement programs, and statewide education reform. 15 percent amounts to approximately $55 million total (i.e., for all 50 states) in FY 2002.

Under 21st Century Community Learning Centers, each state may use up to 3 percent of its grant funds for discretionary activities such as monitoring and evaluation of local activities funded by the program. 3 percent amounts to nearly $30 million total (i.e., for all 50 states) in FY 2002.

**Recommendations to foster the widespread, effective *use* of proven interventions by recipients of federal education funds.**

9. **We strongly support the Department's initiative to develop a "What Works Clearinghouse" – a user-friendly, online database summarizing interventions that have been proven effective** and replicable in scientifically rigorous studies. Such a Clearinghouse could provide many educational agencies and educators, for the first time, with the ability to readily identify and replicate research-proven interventions. We recommend that the Department, in designing and implementing the Clearinghouse, use the randomized controlled trial as the main criterion for deciding what is effective, based on the strong empirical evidence supporting such an approach, discussed earlier.

   While we believe that providing credible, user-friendly information to educational agencies and educators on proven interventions will be an important step forward, the experience of the Comprehensive School Reform Demonstration (CSRD) program, discussed earlier, suggests that it may not be sufficient to spur the widespread use of these interventions. In that example, neither a legislative requirement for the use of evidence-based school reform models, nor the existence of user-friendly, *Consumer Reports*-type ratings of school reform models based on evidence of effectiveness, proved sufficient to spur the widespread use of evidence-based school reforms. We therefore believe that further incentives, as follows, are needed to spur the widespread use of proven interventions by grant recipients.

10. **The Department should give grant recipients significant incentives to implement proven, replicable interventions on a large-scale basis.** Specifically:

    a. **The Department should require grant applicants to provide a plan, in their applications, for the widespread implementation of proven interventions, with quantifiable goals.** This requirement would apply to applicants in both the Department's formula and discretionary grant programs. The requirement would be a logical step toward implementation of the "scientifically-based research" provision in programs affected by that provision; however, we believe the Department should apply it in other programs, not covered by that provision, as well. Of course, before this requirement can be applied in a particular program, there must exist one or more interventions in that program area that has been proven effective and replicable in randomized controlled trials. The best scenario would be the existence of several such interventions in the program area. However, we believe the requirement should be applied even if only one proven intervention exists, because doing so will (i) help to rapidly disseminate that effective intervention, and (ii) provide a powerful incentive for providers of other interventions to carry out randomized trials to establish the effectiveness of their interventions, as well.

       **Each applicant's plan would:**

       i. **Identify the research-proven interventions, drawn from the What Works Clearinghouse or elsewhere, that the applicant plans to implement** in its activities. We believe it is

important that the applicant be allowed to choose which interventions to include in its plan – even if those interventions are not listed in the Clearinghouse – provided the applicant can show that the interventions are backed by randomized trials showing their effectiveness.

ii. **Cite and briefly summarize the randomized controlled trials supporting these interventions,** along with any other evidence of their effectiveness.

iii. **Discuss the applicant's strategy to foster widespread implementation of these interventions** – including the specific protocols that have been proven effective – in the applicant's proposed activities. For example, in the case of a counseling intervention that has been shown effective in reducing truancy, the applicant would discuss its strategy to foster widespread implementation not of counseling to prevent truancy, but of the specific counseling intervention and protocol that has been proven effective.

iv. **Provide measurable, quantitative goals for such widespread implementation** of the proven, replicable interventions (e.g., number of children that will participate in the interventions).

v. **Discuss how the applicant plans to evaluate, after grant award, whether it is meeting its goals for the widespread implementation of these interventions**, and whether it is implementing them with fidelity to the protocols that have been proven effective. As discussed below, in the Department's High Priority Areas, the Department would require an *independent* evaluation of the awardee's implementation of proven interventions.

vi. **Discuss the applicant's track record in implementing proven interventions under its previous grant awards**, or as part of its other activities.

b. **In the Department's discretionary grant programs, this plan would be an important factor in the proposal evaluation process.** Also, in Department programs that provide formula grants to the states, which then make *discretionary* grants to local entities (e.g. Reading First, Comprehensive School Reform), the Department could provide incentives to the states (such as those discussed in recommendation 8) to make the plan an important factor in the state-level proposal evaluation process.

c. **In High Priority Areas, the Department would require an *independent* evaluation of whether the applicant, after grant award, is meeting its goals for the widespread implementation** of these interventions, and whether it is implementing them with fidelity to the protocols that have been proven effective. Specifically, the Department would require grant applicants in these High Priority Areas to team with a Process Evaluator who, after the grant award, would carry out the independent evaluation. To ensure the independence of the evaluation, the Department would provide funding for the evaluation directly to the Process Evaluator, using funds identified per recommendation 4 above.

Process Evaluators, in carrying out these evaluations, will likely need to develop creative ways to leverage their limited resources, given the size of the effort required to verify implementation of the proven interventions at the school or classroom level. For example a Process Evaluator, in assessing the implementation of a school-based intervention provided by an outside vendor (e.g., Reading Recovery, Success for All), may decide to collaborate with the vendor in carrying out the process evaluation. In other evaluations, the Process Evaluator may decide to collaborate with the appropriate state and local educational agencies, in ways that don't compromise the independence of the evaluation. Many state and local educational agencies are themselves

concerned about whether interventions are being implemented with fidelity in the schools and classrooms in their jurisdiction, and may welcome and assist the Process Evaluator in its efforts.

d. **The Department would issue an annual report summarizing the progress of each state agency and other major grantee in implementing research-proven interventions.** The report would be based on the results of the process evaluations discussed above. The Department would issue a high-profile public release of each annual report (e.g., sending a copy to the Governors and Congressional Education Committees), and highlight examples of successful implementation, in order to build momentum for the dissemination of proven interventions.

**Other Recommendation**

11. **The Department should experiment with, and rigorously evaluate, alternative strategies for incorporating an evidence-based approach into its grant-making procedures.** Suppose, for example, the Department wishes to evaluate the impact on the Safe and Drug-Free Schools program of (i) requiring each applicant, in one of the program's discretionary grant competitions, to provide a plan for the widespread implementation of research-proven interventions to prevent substance abuse (per recommendation 10, above), and (ii) after grant award, independently evaluating the implementation of that plan. The Department could carry out a randomized trial to evaluate whether this reformed granting process leads to better outcomes than the current granting process, as follows:

    a. The Department, in the grant application package, would require each potential applicant to send the Department a brief notification of its intent to apply for funding.

    b. The Department would then randomly assign the entities that intend to apply for funding to either the reformed granting process or the current granting process.

    c. The Department would enlist an Impact Evaluator to carry out the evaluation – in our example, to compare the substance-abuse outcomes of students under the two alternative granting processes. Funding for the evaluation would come from sources identified per recommendation 4 above.

    Randomized trials such as these would enable the Department to systematically experiment with, and evaluate, alternative evidence-based reforms in the Department's grant-making processes. We believe such experimentation and evaluation is critical because, as the experience of the CSRD program pointedly demonstrates, there is no way *a priori* to know for certain which of these reforms will prove effective and which will not.

    Structuring the grant-making process as a randomized trial in this way is only possible in the Department's discretionary grant – as opposed to formula grant – programs. Importantly, however, it could also work in Department programs that provide formula grants to the states, which then make discretionary grants to local entities (e.g. Reading First, Comprehensive School Reform) – *if* the Department can persuade one or more states to use its grant-making process to experiment with and evaluate evidence-based reforms in this way.

**Conclusion: These recommendations would have both a near-term and longer-term impact in improving American education.**

We believe that some of the recommendations we have outlined would significantly improve educational outcomes even in the near term. Specifically, in areas where proven, replicable

interventions currently exist – such as early reading instruction and substance-abuse prevention – our recommendations to spur their widespread use could yield major benefits in the next few years, as students participate in these interventions in increasing numbers.

Our recommendations to build the knowledge base of proven, replicable interventions would likely yield benefits over a longer time frame, for the following reason. Once the Department and its grantees begin systematically to test existing and new interventions in randomized trials, we can expect, based on past results of such trials, that many interventions will be found to be ineffective or marginally effective, and a smaller number will be found truly effective, particularly when replicated on a large scale. Both types of findings – negative and positive – are essential in building cumulative knowledge about how to effectively educate children, but the nature of the process, including the likelihood of many misfires, means that it will take time.

Ultimately, we believe that our proposed strategy for implementing the scientifically-based research concept would create a new dynamic for rapid, cumulative advances in the effectiveness of American education. Such advances would be driven by: (i) an ever-expanding knowledge base of educational interventions proven effective in randomized controlled trials, and (ii) an efficient process for translating that knowledge into widespread practice.

The steps we have outlined to create this new dynamic are all designed to be implemented by the Department within its existing statutory authority, and require no new action from Congress. However, they will require from the Department a major, sustained commitment of resources and effort over time. For such an effort to succeed, we believe the Department's leadership must embrace it as a central priority.

## Attachment 1 – Department of Education Press Release Announcing the Collaborative Initiative With the Coalition for Evidence-Based Policy

U.S. Department of Education
Office of Public Affairs, News Branch
400 Maryland Ave., S.W.
Washington, D.C. 20202

FOR RELEASE
April 24, 2002
Contact: Public Affairs
David Thomas (202) 401-1579

Valerie Reyna (202) 219-1385
(Non-media inquiries)

**EDUCATION DEPARTMENT, COALITION FOR EVIDENCE-BASED POLICY TO COLLABORATE ON NEW INITIATIVE**

The U.S. Department of Education today announced that it will work with the Council for Excellence in Government's Coalition for Evidence-Based Policy on a new initiative to explore how the department can most effectively advance evidence-based approaches to federal education policy.

This initiative is designed to help the department achieve the goal of transforming education into an evidence-based field as outlined in the department's recently released strategic plan. The initiative will also help advance the key principle in the No Child Left Behind Act of 2001: that federal funds should support programs and strategies that are based on scientifically based research.

"The success of federal education programs depends ultimately on the ability to evaluate which programs are working as intended and which are not. With rigorous evidence on effectiveness we can begin to focus our resources on programs that work," said Grover "Russ" Whitehurst, assistant secretary for educational research and improvement. "The coalition has brought together representatives from various disciplines and areas of policy to provide an independent analysis of how the department can use its authority strategically and effectively to embed the collection and use of evidence in all of its programs. This is a very important initiative, and I look forward to working with the coalition."

The Coalition for Evidence-Based Policy is sponsored by the nonprofit, nonpartisan Council for Excellence in Government. Its mission is to promote government policymaking based on rigorous evidence of program effectiveness. This coalition initiative will specifically address how the department can most effectively use its statutory and policy authority to (1) promote rigorous evaluation of education programs and strategies, and (2) ensure that program funds support research-proven educational activities. The initiative will draw on successful precedents for evidence-based approaches from medicine and welfare policy, where rigorous evaluation has played a prominent role in policy and funding decisions.

More information on the Coalition for Evidence-Based Policy and its activities can be found at http://www.excelgov.org/performance/evidence/execsumm.htm.

## Attachment 2 – Background on the Coalition For Evidence-Based Policy

**The Coalition for Evidence-Based Policy was launched in fall 2001 under the sponsorship of the Council for Excellence in Government**, a nonpartisan, nonprofit organization, founded in 1983, that works to improve the performance of government at all levels and government's place in the lives and esteem of American citizens.

**The Coalition's mission is to promote government policymaking based on rigorous evidence of program effectiveness.** Its primary agenda is to work with key policymakers in the federal agencies and Congress to incorporate an "evidence-based approach," as follows, into social and economic programs:

1. Allocation of government resources -- *Government funding or other benefits should be allocated to activities that either (i) have been shown to be effective through rigorous empirical studies, or (ii) as a condition of funding (or other benefit), will undergo an independent, rigorous evaluation.*

2. Government R&D investment -- *The government should undertake, or otherwise advance, a systematic R&D effort to identify and/or develop effective government interventions. This R&D effort should use rigorous study designs, such as the randomized controlled trial.*

**The Coalition's Board of Advisors includes distinguished, evidence-based policymakers and scholars from a number of different fields who together comprise a bipartisan balance.** What follows is a list of the Coalition's Board members and their brief bios:

- Robert Boruch, professor of education and statistics at the University of Pennsylvania, co-director of the Center for Research and Evaluation of Social Policy, co-chair of the International Campbell Collaboration on Systematic Reviews, member of Advisory Committees to DoEd and GAO, author of *Randomized Experiments for Planning and Evaluations*.

- Jonathan Crane, senior fellow at the Progressive Policy Institute, former senior domestic policy advisor to Vice President Gore, adjunct professor of public policy at the University of Illinois, editor of *Social Programs That Work*.

- David Ellwood, professor of political economy and former academic dean at the Kennedy School of Government at Harvard University, former Assistant Secretary of HHS for Planning and Evaluation and co-chair of President Clinton's welfare reform effort, author of numerous books and articles including *Welfare Realities: From Rhetoric to Reform* (with Mary Jo Bane).

- Judith Gueron, president of Manpower Demonstration Research Corporation (MDRC); principal researcher on numerous MDRC random-assignment studies in areas including welfare, employment, and poverty reduction; member of the National Academy of Sciences Committee on Poverty and Family Assistance; author of *From Welfare To Work* (with Edward Pauly).

- Ron Haskins, senior fellow at the Brookings Institution and co-director of the Welfare Reform & Beyond initiative at Brookings, former majority (Republican) staff director of the House Ways and Means Subcommittee on Human Resources, former research professor at UNC-Chapel Hill, author of numerous articles on welfare reform and other areas of social policy.

- Robert Hoyt, COO of Market Direct, a construction procurement company, and Managing Director of Alignware, a company that orchestrates reverse electronic auctions; former assistant professor of psychology at Mount Sinai Medical School in New York, he left a clinical practice to assume control of a small industrial distribution company, which he took public in 1998.

- Blair Hull, founder and former CEO, Hull Trading Company, one of the world's premier market making trading firms; pioneer of the transition from open outcry trading to automated trading; current chairman and CEO, Matlock Capital LLC.

- David Kessler, Dean of the Yale University School of Medicine, former Commissioner of the FDA (appointed by President Bush and reappointed by President Clinton), former medical director of the Hospital of the Albert Einstein College of Medicine, author of *A Question of Intent* as well as numerous medical journal articles.

- Diane Ravitch, distinguished visiting fellow at the Hoover Institution, senior fellow at the Brookings Institution, research professor at NYU, member of the National Assessment Governing Board, former Assistant Secretary of Education for Educational Research and Improvement (Bush administration), author of numerous books and articles on education policy and the history of education.

- Laurie Robinson, distinguished senior scholar at the University of Pennsylvania's Jerry Lee Center of Criminology; former Assistant Attorney General in the Department of Justice, heading the Office of Justice Programs (including National Institute of Justice, Bureau of Justice Statistics, and Bureau of Justice Assistance); former director of the ABA's Section on Criminal Justice.

- Isabel Sawhill, senior fellow at the Brookings Institution and co-director of the Welfare Reform & Beyond initiative at Brookings, former Associate Director of OMB, author of numerous books and articles including *Updating America's Social Contract: Economic Growth and Opportunity in the New Century*.

- Martin Seligman, professor of psychology at the University of Pennsylvania, past president of the American Psychological Association, author of more than 150 articles and 15 books, including the best-seller *Learned Optimism*.

- Robert Slavin, co-director of the Center for Research on the Education of Students Placed at Risk at Johns Hopkins University, chairman of the Success for All Foundation, author of more than 200 articles and 15 books including *Educational Psychology: Theory into Practice*.

- Robert Solow, Institute Professor Emeritus at the MIT, recipient of the 1987 Nobel Prize in Economic Science for his seminal contributions to the theory of capital and economic growth, former president of the American Economic Association and the Econometric Society, author of numerous books and articles on the sources of economic growth, the nature of the labor market, and other topics.

- Nicholas Zill, vice president of Westat, Inc., a survey research firm; senior advisor to DoEd on the Early Childhood Longitudinal Study; founder and former executive director of Child Trends, a nonprofit organization working to improve policy-relevant data on children; author of numerous articles on family behavior and its effects on children.

## Attachment 3 – The School Reform Models Adopted By Schools Participating in the Comprehensive School Reform Demonstration (CSRD) Program

The following table lists the various school reform models, the American Institutes for Research (AIR) rating of the evidence base for each model, and the percentage of CSRD-funded schools adopting each model in 1998 and 2001-2002. As shown in the table, the percentage of CSRD-funded schools adopting models with a "strong" or "promising" evidence rating has decreased over time.

| AIR Rating | All States 1998 (n=1855) | Survey of 19 States 2001-02 (n=560) | Change |
|---|---|---|---|
| **Strong** | | | |
| Success for All/Roots&Wings | 16.5 | 5.7 | |
| Direct Instruction | 2.7 | 1.1 | |
| High Schools That Work | 2.8 | 5.2 | |
| | 22.0 | 12.0 | -10.0% |
| **Promising** | | | |
| Community for Learning | 0.8 | 2.3 | |
| Core Knowledge | 2.2 | 2.0 | |
| Different Ways of Knowing | 1.5 | 0.5 | |
| Expeditionary Learning | 0.8 | 0.5 | |
| School Development Program | 1.3 | 1.1 | |
| | 6.6 | 6.4 | -0.2% |
| **Marginal** | | | |
| Accelerated Schools | 5.2 | 7.7 | |
| Onward to Excellence | 0.4 | 2.3 | |
| | 5.6 | 10.0 | +4.4% |
| **Mixed or No Research** | | | |
| America's Choice | 2.1 | 0 | |
| ATLAS | 0.4 | 1.4 | |
| Coalition of Essential Schools | 1.2 | 1.8 | |
| Co-NECT | 1.4 | 2.3 | |
| Modern Red Schoolhouse | 0.8 | 0.4 | |
| | 5.9 | 5.9 | 0 |
| **Not on AIR List** (receiving 1% or more of 01-02 grants) | | | |
| AVID | 1.4 | 3.0 | |
| CCC | 0.7 | 1.4 | |
| Effective Schools | 1.4 | 3.0 | |
| Literacy in Action | -- | 2.0 | |
| Literacy Collaborative | 0.7 | 2.9 | |
| HOSTS | 1.8 | 2.0 | |
| Lightspan | 6.0 | 3.4 | |
| Reading Recovery | 1.4 | 1.1 | |
| Target Teach | -- | 1.8 | |
| Understanding by Design | -- | 1.3 | |
| School Renaissance | -- | 1.3 | |
| | 12.0 | 23.2 | +11.2% |
| All Others | 47.9 | 42.5 | |
| Total | 100% | 100% | |

*U.S. Department of Education*
*Institute of Education Sciences (IES)*
*National Library of Education (NLE)*
*Educational Resources Information Center (ERIC)*

# Reproduction Release

(Specific Document)

## I. DOCUMENT IDENTIFICATION:

| | |
|---|---|
| Title: Bringing Evidence-Driven Progress to Education: A Recommended Strategy for the U.S. Department of Education | |
| Author(s): Coalition for Evidence-Based Policy | |
| Corporate Source: Council for Excellence in Government | Publication Date: November 18, 2002 |

## II. REPRODUCTION RELEASE:

In order to disseminate as widely as possible timely and significant materials of interest to the educational community, documents announced in the monthly abstract journal of the ERIC system, Resources in Education (RIE), are usually made available to users in microfiche, reproduced paper copy, and electronic media, and sold through the ERIC Document Reproduction Service (EDRS). Credit is given to the source of each document, and, if reproduction release is granted, one of the following notices is affixed to the document.

If permission is granted to reproduce and disseminate the identified document, please CHECK ONE of the following three options and sign in the indicated space following.

| The sample sticker shown below will be affixed to all Level 1 documents | The sample sticker shown below will be affixed to all Level 2A documents | The sample sticker shown below will be affixed to all Level 2B documents |
|---|---|---|
| PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL HAS BEEN GRANTED BY SAMPLE TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC) | PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL IN MICROFICHE, AND IN ELECTRONIC MEDIA FOR ERIC COLLECTION SUBSCRIBERS ONLY, HAS BEEN GRANTED BY SAMPLE TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC) | PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL IN MICROFICHE ONLY HAS BEEN GRANTED BY SAMPLE TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC) |
| Level 1 | Level 2A | Level 2B |
| ↑ ☑ | ↑ ☐ | ↑ ☐ |
| Check here for Level 1 release, permitting reproduction and dissemination in microfiche or other ERIC archival media (e.g. electronic) *and* paper copy. | Check here for Level 2A release, permitting reproduction and dissemination in microfiche and in electronic media for ERIC archival collection subscribers only | Check here for Level 2B release, permitting reproductic and dissemination in microfiche only |

Documents will be processed as indicated provided reproduction quality permits.

If permission to reproduce is granted, but no box is checked, documents will be processed at Level 1.

*I hereby grant to the Educational Resources Information Center (ERIC) nonexclusive permission to reproduce and disseminc document as indicated above. Reproduction from the ERIC microfiche, or electronic media by persons other than ERIC empi system contractors requires permission from the copyright holder. Exception is made for non-profit reproduction by librari service agencies to satisfy information needs of educators in response to discrete inquiries.*

| Signature: | Printed Name/Position/Title: | |
|---|---|---|
| | Jon Baron / Executive Director / Coalition for Evidence-Based Policy | |
| Organization/Address: 1301 K Street, N.W., Suite 450 West Washington, DC 20005 | Telephone: 202-530-3279 | Fax: 202-728-0422 |
| | E-mail Address: jbaron@excelgov.org | Date: February 4, 2003 |

## III. DOCUMENT AVAILABILITY INFORMATION (FROM NON-ERIC SOURCE):

If permission to reproduce is not granted to ERIC, or, if you wish ERIC to cite the availability of the document from another source, please provide the following information regarding the availability of the document. (ERIC will not announce a document unless it is publicly available, and a dependable source can be specified. Contributors should also be aware that ERIC selection criteria are significantly more stringent for documents that cannot be made available through EDRS.)

| |
|---|
| Publisher/Distributor: |
| Address: |
| Price: |

## IV. REFERRAL OF ERIC TO COPYRIGHT/REPRODUCTION RIGHTS HOLDER:

If the right to grant this reproduction release is held by someone other than the addressee, please provide the appropriate name and address:

| |
|---|
| Name: |
| Address: |

## V. WHERE TO SEND THIS FORM:

Send this form to the following ERIC Clearinghouse:

Document Acquisitions Department
ERIC Clearinghouse on Educational Management
5207 University of Oregon
Eugene, Oregon 97403-5207